

INU

7-28-02



ENTERED

PCT09

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/807,802A

DATE: 04/29/2002

TIME: 11:21:58

Input Set : A:\GNVPN.031.txt

Output Set: N:\CRF3\04292002\I807802A.raw

3 <110> APPLICANT: Wilson, James M.
 4 Xiao, Weidong
 6 <120> TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
 7 Vectors and Host Cells Containing Same
 9 <130> FILE REFERENCE: GNVPN.031USA
 11 <140> CURRENT APPLICATION NUMBER: US 09/807,802A
 C--> 12 <141> CURRENT FILING DATE: 2002-02-21
 14 <150> PRIOR APPLICATION NUMBER: US 60/107,114
 15 <151> PRIOR FILING DATE: 1998-11-05
 17 <150> PRIOR APPLICATION NUMBER: PCT/US99/25694
 18 <151> PRIOR FILING DATE: 1999-11-02
 20 <160> NUMBER OF SEQ ID NOS: 20
 22 <170> SOFTWARE: PatentIn version 3.1
 24 <210> SEQ ID NO: 1
 25 <211> LENGTH: 4718
 26 <212> TYPE: DNA
 27 <213> ORGANISM: AAV-1
 29 <220> FEATURE:
 30 <221> NAME/KEY: CDS
 31 <222> LOCATION: (335)..(2206)
 32 <223> OTHER INFORMATION:
 34 <220> FEATURE:
 35 <221> NAME/KEY: CDS
 36 <222> LOCATION: (2223)..(4430)
 37 <223> OTHER INFORMATION:
 39 <400> SEQUENCE: 1
 40 ttgcccactc cctctctgcg cgctcgctcg ctcggtgggg cctgcggacc aaaggtccgc 60
 42 agacggcaga gctctgctct gccggcccca ccgagcgagc gagcgcgag agagggagtg 120
 44 ggcaactcca tctactagggg taatgcgcga gcgcctccca cgctgccgcg tcagcgctga 180
 46 cgtaaattac gtcatagggg agtggtcctg tattagctgt cacgtgagtg cttttgagac 240
 48 attttgagac accacgtggc catttagggg atatatggcc gagtgcgcga gcaggatctc 300
 50 cattttgacc gcgaaatttg aacgagcgagc agcc atg ccg ggc ttc tac gag atc 355
 51 Met Pro Gly Phe Tyr Glu Ile
 52 1 5
 54 gtg atc aag gtg ccg agc gac ctg gac gag cac ctg ccg ggc att tct 403
 55 Val Ile Lys Val Pro Ser Asp Leu Asp Glu His Leu Pro Gly Ile Ser
 56 10 15 20
 58 gac tcg ttt gtg agc tgg gtg gcc gag aag gaa tgg gag ctg ccc ccg 451
 59 Asp Ser Phe Val Ser Trp Val Ala Glu Lys Glu Trp Glu Leu Pro Pro
 60 25 30 35
 62 gat tct gac atg gat ctg aat ctg att gag cag gca ccc ctg acc gtg 499
 63 Asp Ser Asp Met Asp Leu Asn Leu Ile Glu Gln Ala Pro Leu Thr Val
 64 40 45 50 55

RAW SEQUENCE LISTING

DATE: 04/29/2002

PATENT APPLICATION: US/09/807,802A

TIME: 11:21:58

Input Set : A:\GNVPM.031.txt

Output Set: N:\CRF3\04292002\I807802A.raw

66	gcc	gag	aag	ctg	cag	cgc	gac	ttc	ctg	gtc	caa	tgg	cgc	cgc	gtg	agt	547
67	Ala	Glu	Lys	Leu	Gln	Arg	Asp	Phe	Leu	Val	Gln	Trp	Arg	Arg	Val	Ser	
68				60						65					70		
70	aag	gcc	ccg	gag	gcc	ctc	ttc	ttt	gtt	cag	ttc	gag	aag	ggc	gag	tcc	595
71	Lys	Ala	Pro	Glu	Ala	Leu	Phe	Phe	Val	Gln	Phe	Glu	Lys	Gly	Glu	Ser	
72				75					80					85			
74	tac	ttc	cac	ctc	cat	att	ctg	gtg	gag	acc	acg	ggg	gtc	aaa	tcc	atg	643
75	Tyr	Phe	His	Leu	His	Ile	Leu	Val	Glu	Thr	Thr	Gly	Val	Lys	Ser	Met	
76			90					95					100				
78	gtg	ctg	ggc	cgc	ttc	ctg	agt	cag	att	agg	gac	aag	ctg	gtg	cag	acc	691
79	Val	Leu	Gly	Arg	Phe	Leu	Ser	Gln	Ile	Arg	Asp	Lys	Leu	Val	Gln	Thr	
80		105					110					115					
82	atc	tac	cgc	ggg	atc	gag	ccg	acc	ctg	ccc	aac	tgg	ttc	gcg	gtg	acc	739
83	Ile	Tyr	Arg	Gly	Ile	Glu	Pro	Thr	Leu	Pro	Asn	Trp	Phe	Ala	Val	Thr	
84	120				125						130				135		
86	aag	acg	cgt	aat	ggc	gcc	gga	ggg	ggg	aac	aag	gtg	gtg	gac	gag	tgc	787
87	Lys	Thr	Arg	Asn	Gly	Ala	Gly	Gly	Gly	Asn	Lys	Val	Val	Asp	Glu	Cys	
88				140					145					150			
90	tac	atc	ccc	aac	tac	ctc	ctg	ccc	aag	act	cag	ccc	gag	ctg	cag	tgg	835
91	Tyr	Ile	Pro	Asn	Tyr	Leu	Leu	Pro	Lys	Thr	Gln	Pro	Glu	Leu	Gln	Trp	
92			155					160					165				
94	gcg	tgg	act	aac	atg	gag	gag	tat	ata	agc	gcc	tgt	ttg	aac	ctg	gcc	883
95	Ala	Trp	Thr	Asn	Met	Glu	Glu	Tyr	Ile	Ser	Ala	Cys	Leu	Asn	Leu	Ala	
96			170					175					180				
98	gag	cgc	aaa	cgg	ctc	gtg	gcg	cag	cac	ctg	acc	cac	gtc	agc	cag	acc	931
99	Glu	Arg	Lys	Arg	Leu	Val	Ala	Gln	His	Leu	Thr	His	Val	Ser	Gln	Thr	
100		185					190					195					
102	cag	gag	cag	aac	aag	gag	aat	ctg	aac	ccc	aat	tct	gac	gcg	cct	gtc	979
103	Gln	Glu	Gln	Asn	Lys	Glu	Asn	Leu	Asn	Pro	Asn	Ser	Asp	Ala	Pro	Val	
104	200				205					210				215			
106	atc	cgg	tca	aaa	acc	tcc	gcg	cgc	tac	atg	gag	ctg	gtc	ggg	tgg	ctg	1027
107	Ile	Arg	Ser	Lys	Thr	Ser	Ala	Arg	Tyr	Met	Glu	Leu	Val	Gly	Trp	Leu	
108				220					225					230			
110	gtg	gac	cgg	ggc	atc	acc	tcc	gag	aag	cag	tgg	atc	cag	gag	gac	cag	1075
111	Val	Asp	Arg	Gly	Ile	Thr	Ser	Glu	Lys	Gln	Trp	Ile	Gln	Glu	Asp	Gln	
112			235					240					245				
114	gcc	tcg	tac	atc	tcc	ttc	aac	gcc	gct	tcc	aac	tcg	cgg	tcc	cag	atc	1123
115	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala	Ser	Asn	Ser	Arg	Ser	Gln	Ile	
116			250					255					260				
118	aag	gcc	gct	ctg	gac	aat	gcc	ggc	aag	atc	atg	gcg	ctg	acc	aaa	tcc	1171
119	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys	Ile	Met	Ala	Leu	Thr	Lys	Ser	
120		265				270				275							
122	gcg	ccc	gac	tac	ctg	gta	ggc	ccc	gct	ccg	ccc	gcg	gac	att	aaa	acc	1219
123	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Pro	Ala	Pro	Pro	Ala	Asp	Ile	Lys	Thr	
124	280				285					290				295			
126	aac	cgc	atc	tac	cgc	atc	ctg	gag	ctg	aac	ggc	tac	gaa	cct	gcc	tac	1267
127	Asn	Arg	Ile	Tyr	Arg	Ile	Leu	Glu	Leu	Asn	Gly	Tyr	Glu	Pro	Ala	Tyr	
128				300					305				310				
130	gcc	ggc	tcc	gtc	ttt	ctc	ggc	tgg	gcc	cag	aaa	agg	ttc	ggg	aag	cgc	1315

RAW SEQUENCE LISTING

DATE: 04/29/2002

PATENT APPLICATION: US/09/807,802A

TIME: 11:21:58

Input Set : A:\GNVFN.031.txt

Output Set: N:\CRF3\04292002\I807802A.raw

131	Ala	Gly	Ser	Val	Phe	Leu	Gly	Trp	Ala	Gln	Lys	Arg	Phe	Gly	Lys	Arg	
132				315					320					325			
134	aac	acc	atc	tgg	ctg	ttt	ggg	ccg	gcc	acc	acg	ggc	aag	acc	aac	atc	1363
135	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala	Thr	Thr	Gly	Lys	Thr	Asn	Ile	
136			330					335					340				
138	gcg	gaa	gcc	atc	gcc	cac	gcc	gtg	ccc	ttc	tac	ggc	tgc	gtc	aac	tgg	1411
139	Ala	Glu	Ala	Ile	Ala	His	Ala	Val	Pro	Phe	Tyr	Gly	Cys	Val	Asn	Trp	
140		345					350					355					
142	acc	aat	gag	aac	ttt	ccc	ttc	aat	gat	tgc	gtc	gac	aag	atg	gtg	atc	1459
143	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp	Cys	Val	Asp	Lys	Met	Val	Ile	
144	360					365					370				375		
146	tgg	tgg	gag	gag	ggc	aag	atg	acg	gcc	aag	gtc	gtg	gag	tcc	gcc	aag	1507
147	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala	Lys	Val	Val	Glu	Ser	Ala	Lys	
148				380						385					390		
150	gcc	att	ctc	ggc	ggc	agc	aag	gtg	cgc	gtg	gac	caa	aag	tgc	aag	tcg	1555
151	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg	Val	Asp	Gln	Lys	Cys	Lys	Ser	
152			395					400					405				
154	tcc	gcc	cag	atc	gac	ccc	acc	ccc	gtg	atc	gtc	acc	tcc	aac	acc	aac	1603
155	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val	Ile	Val	Thr	Ser	Asn	Thr	Asn	
156		410					415					420					
158	atg	tgc	gcc	gtg	att	gac	ggg	aac	agc	acc	acc	ttc	gag	cac	cag	cag	1651
159	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser	Thr	Thr	Phe	Glu	His	Gln	Gln	
160		425				430					435						
162	ccg	ttg	cag	gac	cgg	atg	ttc	aaa	ttt	gaa	ctc	acc	cgc	cgt	ctg	gag	1699
163	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe	Glu	Leu	Thr	Arg	Arg	Leu	Glu	
164	440				445					450					455		
166	cat	gac	ttt	ggc	aag	gtg	aca	aag	cag	gaa	gtc	aaa	gag	ttc	ttc	cgc	1747
167	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln	Glu	Val	Lys	Glu	Phe	Phe	Arg	
168			460					465					470				
170	tgg	gcg	cag	gat	cac	gtg	acc	gag	gtg	gcg	cat	gag	ttc	tac	gtc	aga	1795
171	Trp	Ala	Gln	Asp	His	Val	Thr	Glu	Val	Ala	His	Glu	Phe	Tyr	Val	Arg	
172			475				480					485					
174	aag	ggt	gga	gcc	aac	aaa	aga	ccc	gcc	ccc	gat	gac	gcg	gat	aaa	agc	1843
175	Lys	Gly	Gly	Ala	Asn	Lys	Arg	Pro	Ala	Pro	Asp	Asp	Ala	Asp	Lys	Ser	
176		490					495					500					
178	gag	ccc	aag	cgg	gcc	tgc	ccc	tca	gtc	gcg	gat	cca	tcg	acg	tca	gac	1891
179	Glu	Pro	Lys	Arg	Ala	Cys	Pro	Ser	Val	Ala	Asp	Pro	Ser	Thr	Ser	Asp	
180		505				510					515						
182	gcg	gaa	gga	gct	ccg	gtg	gac	ttt	gcc	gac	agg	tac	caa	aac	aaa	tgt	1939
183	Ala	Glu	Gly	Ala	Pro	Val	Asp	Phe	Ala	Asp	Arg	Tyr	Gln	Asn	Lys	Cys	
184	520				525					530					535		
186	tct	cgt	cac	gcg	ggc	atg	ctt	cag	atg	ctg	ttt	ccc	tgc	aag	aca	tgc	1987
187	Ser	Arg	His	Ala	Gly	Met	Leu	Gln	Met	Leu	Phe	Pro	Cys	Lys	Thr	Cys	
188			540				545					550					
190	gag	aga	atg	aat	cag	aat	ttc	aac	att	tgc	ttc	acg	cac	ggg	acg	aga	2035
191	Glu	Arg	Met	Asn	Gln	Asn	Phe	Asn	Ile	Cys	Phe	Thr	His	Gly	Thr	Arg	
192			555				560					565					
194	gac	tgt	tca	gag	tgc	ttc	ccc	ggc	gtg	tca	gaa	tct	caa	ccg	gtc	gtc	2083
195	Asp	Cys	Ser	Glu	Cys	Phe	Pro	Gly	Val	Ser	Glu	Ser	Gln	Pro	Val	Val	

RAW SEQUENCE LISTING

DATE: 04/29/2002

PATENT APPLICATION: US/09/807,802A

TIME: 11:21:58

Input Set : A:\GNVPM.031.txt

Output Set: N:\CRF3\04292002\I807802A.raw

196	570	575	580	
198	aga aag agg acg tat cgg aaa ctc tgt gcc att cat cat ctg ctg ggg	2131		
199	Arg Lys Arg Thr Tyr Arg Lys Leu Cys Ala Ile His His Leu Leu Gly			
200	585	590	595	
202	cgg gct ccc gag att gct tgc tcg gcc tgc gat ctg gtc aac gtg gac	2179		
203	Arg Ala Pro Glu Ile Ala Cys Ser Ala Cys Asp Leu Val Asn Val Asp			
204	600	605	610	615
206	ctg gat gac tgt gtt tct gag caa taa atgacttaaa ccaggt atg gct gcc	2231		
207	Leu Asp Asp Cys Val Ser Glu Gln		Met Ala Ala	
208	620		625	
210	gat ggt tat ctt cca gat tgg ctc gag gac aac ctc tct gag ggc att	2279		
211	Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser Glu Gly Ile			
212	630	635	640	
214	cgc gag tgg tgg gac ttg aaa cct gga gcc ccg aag ccc aaa gcc aac	2327		
215	Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro Lys Ala Asn			
216	645	650	655	
218	cag caa aag cag gac gac ggc cgg ggt ctg gtg ctt cct ggc tac aag	2375		
219	Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro Gly Tyr Lys			
220	660	665	670	
222	tac ctc gga ccc ttc aac gga ctc gac aag ggg gag ccc gtc aac gcg	2423		
223	Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro Val Asn Ala			
224	675	680	685	690
226	gcg gac gca gcg gcc ctc gag cac gac aag gcc tac gac cag cag ctc	2471		
227	Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp Gln Gln Leu			
228	695	700	705	
230	aaa gcg ggt gac aat ccg tac ctg cgg tat aac cac gcc gac gcc gag	2519		
231	Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala Asp Ala Glu			
232	710	715	720	
234	ttt cag gag cgt ctg caa gaa gat acg tct ttt ggg ggc aac ctc ggg	2567		
235	Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly Asn Leu Gly			
236	725	730	735	
238	cga gca gtc ttc cag gcc aag aag cgg gtt ctc gaa cct ctc ggt ctg	2615		
239	Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro Leu Gly Leu			
240	740	745	750	
242	gtt gag gaa ggc gct aag acg gct cct gga aag aaa cgt ccg gta gag	2663		
243	Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg Pro Val Glu			
244	755	760	765	770
246	cag tcg cca caa gag cca gac tcc tcc tcg ggc atc ggc aag aca ggc	2711		
247	Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Ile Gly Lys Thr Gly			
248	775	780	785	
250	cag cag ccc gct aaa aag aga ctc aat ttt ggt cag act ggc gac tca	2759		
251	Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr Gly Asp Ser			
252	790	795	800	
254	gag tca gtc ccc gat cca caa cct ctc gga gaa cct cca gca acc ccc	2807		
255	Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro Ala Thr Pro			
256	805	810	815	
258	gct gct gtg gga cct act aca atg gct tca ggc ggt ggc gca cca atg	2855		
259	Ala Ala Val Gly Pro Thr Thr Met Ala Ser Gly Gly Gly Ala Pro Met			
260	820	825	830	

RAW SEQUENCE LISTING

DATE: 04/29/2002

PATENT APPLICATION: US/09/807,802A

TIME: 11:21:58

Input Set : A:\GNVPN.031.txt

Output Set: N:\CRF3\04292002\I807802A.raw

262	gca gac aat aac gaa ggc gcc gac gga gtg ggt aat gcc tca gga aat	2903
263	Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala Ser Gly Asn	
264	835 840 845 850	
266	tgg cat tgc gat tcc aca tgg ctg ggc gac aga gtc atc acc acc agc	2951
267	Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile Thr Thr Ser	
268	855 860 865	
270	acc cgc acc tgg gcc ttg ccc acc tac aat aac cac ctc tac aag caa	2999
271	Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln	
272	870 875 880	
274	atc tcc agt gct tca acg ggg gcc agc aac gac aac cac tac ttc ggc	3047
275	Ile Ser Ser Ala Ser Thr Gly Ala Ser Asn Asp Asn His Tyr Phe Gly	
276	885 890 895	
278	tac agc acc ccc tgg ggg tat ttt gat ttc aac aga ttc cac tgc cac	3095
279	Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His	
280	900 905 910	
282	ttt tca cca cgt gac tgg cag cga ctc atc aac aac aat tgg gga ttc	3143
283	Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Trp Gly Phe	
284	915 920 925 930	
286	cgg ccc aag aga ctc aac ttc aaa ctc ttc aac atc caa gtc aag gag	3191
287	Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val Lys Glu	
288	935 940 945	
290	gtc acg acg aat gat ggc gtc aca acc atc gct aat aac ctt acc agc	3239
291	Val Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn Leu Thr Ser	
292	950 955 960	
294	acg gtt caa gtc ttc tcg gac tcg gag tac cag ctt ccg tac gtc ctc	3287
295	Thr Val Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro Tyr Val Leu	
296	965 970 975	
298	ggc tct gcg cac cag ggc tgc ctc cct ccg ttc ccg gcg gac gtg ttc	3335
299	Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe	
300	980 985 990	
302	atg att ccg caa tac ggc tac ctg acg ctc aac aat ggc agc caa	3380
303	Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser Gln	
304	995 1000 1005	
306	gcc gtg gga cgt tca tcc ttt tac tgc ctg gaa tat ttc cct tct	3425
307	Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser	
308	1010 1015 1020	
310	cag atg ctg aga acg ggc aac aac ttt acc ttc agc tac acc ttt	3470
311	Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe	
312	1025 1030 1035	
314	gag gaa gtg cct ttc cac agc agc tac gcg cac agc cag agc ctg	3515
315	Glu Glu Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu	
316	1040 1045 1050	
318	gac cgg ctg atg aat cct ctc atc gac caa tac ctg tat tac ctg	3560
319	Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu	
320	1055 1060 1065	
322	aac aga act caa aat cag tcc gga agt gcc caa aac aag gac ttg	3605
323	Asn Arg Thr Gln Asn Gln Ser Gly Ser Ala Gln Asn Lys Asp Leu	
324	1070 1075 1080	
326	ctg ttt agc cgt ggg tct cca gct ggc atg tct gtt cag ccc aaa	3650

VERIFICATION SUMMARY

DATE: 04/29/2002

PATENT APPLICATION: US/09/807,802A

TIME: 11:21:59

Input Set : A:\GNVPN.031.txt

Output Set: N:\CRF3\04292002\I807802A.raw

L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date